

Claims

What is claimed is:

1. A method for redistributing data in a relational data base management system, comprising:  
allocating a buffer associated with a transmitting processing module, the transmitting processing  
module having access to a program, the program capable of managing a redistribution of one or more  
rows associated with one or more database tables;  
comparing the allocated buffer to a portion of the buffer to be occupied by the one or more rows;  
if the allocated buffer is larger than the portion of the buffer to be occupied by the one or more  
rows:
  - 10 storing one or more rows of a database table in the allocated buffer;
  - communicating a message to one or more destination processing modules, the message  
comprising at least some of the one or more rows stored in the allocated buffer;
  - otherwise:
    - executing a many-rows method to redistribute the one or more rows.
- 15 2. The method of Claim 1, wherein the message comprises all of the one or more rows stored  
in the allocated buffer.
3. The method of Claim 1, wherein the transmitting processing module comprises one of a  
20 plurality of processing modules associated with a relational database system.
4. The method of Claim 1, wherein the message is communicated to each of a plurality of  
destination processing elements.
- 25 5. The method of Claim 1, wherein the allocated buffer is capable of storing no more than ten  
(10) rows.
6. The method of Claim 1, wherein the many-rows method comprises:  
communicating from one or more transmitting processing modules a first signal to a plurality of  
30 processing modules within a relational database system, the first signal operable to initiate a row  
receiver task on each of the processing modules;

communicating from one or more of the processing modules a ready-to-receive signal to the one or more transmitting processing modules;

communicating from the one or more transmitting processing modules a second signal comprising the one or more rows associated with the database table;

5 after communication of the last row associated with the database table, communicating from the one or more transmitting processing modules an end-of-data signal to each of the plurality of processing modules.

7. The method of Claim 1, further comprising invoking the program on a single transmitting  
10 processing module.

8. The method of Claim 1, further comprising receiving at each of a plurality of destination processing elements a substantially similar set of the one or more rows stored in the allocated buffer.

15 9. The method of Claim 1, further comprising determining a number of rows to store in the allocated buffer.

20

10. A method for redistributing data in a relational data base management system, comprising:  
invoking a program on one or more of a plurality of transmitting modules, the program capable  
of managing a redistribution of one or more rows associated with one or more database tables;

5 if the program was invoked on a single transmitting module;

executing a few-rows redistribution method to redistribute the one or more rows;

otherwise;

executing a many-rows redistribution method to redistribute the one or more rows.

11. The method of Claim 10, wherein the few-rows row redistribution method comprises:

10 allocating a buffer associated with a transmitting processing module, the transmitting processing  
module having access to the program, the program associated with a single transmitting module;

comparing the allocated buffer to a portion of the buffer to be occupied by the one or more rows;

if the allocated buffer is larger than the portion of the buffer to be occupied by the one or more  
rows:

15 storing one or more rows of a database table in the allocated buffer;

communicating a message to one or more destination modules, the message comprising at  
least some of the one or more rows stored in the allocated buffer;

otherwise:

executing a many-rows method to redistribute the one or more rows.

20

12. The method of Claim 11, wherein the message comprises all of the one or more rows stored  
in the allocated buffer.

25 13. The method of Claim 11, wherein the transmitting module comprises one of a plurality of  
processing modules associated with a relational database system.

14. The method of Claim 11, wherein the message is communicated to each of a plurality of  
destination modules.

30 15. The method of Claim 11, wherein the few-rows row redistribution method further  
comprises determining a number of rows to store in the allocated buffer.

16. The method of Claim 10, wherein the many-rows row redistribution method comprises:  
communicating from one or more transmitting modules a first signal to a plurality of processing  
modules within a relational database system, the first signal operable to initiate a row receiver task on  
each of the processing modules;

5           communicating from one or more of the processing modules a ready-to-receive signal to the one  
or more transmitting modules;

10          communicating from the one or more transmitting modules a second signal comprising the one or  
more rows associated with the database table;

15          after communication of the last row associated with the database table, communicating from the  
one or more transmitting modules an end-of-data signal to each of the plurality of processing modules.

15

18. A computer-readable medium containing computer-executable code for instructing a computer to:

allocate a buffer associated with a transmitting processing module, the transmitting processing module having access to a program, the program capable of managing a redistribution of one or more rows associated with one or more database tables;

5 compare the allocated buffer to a portion of the buffer to be occupied by the one or more rows;  
if the allocated buffer is larger than the portion of the buffer to be occupied by the one or more rows:

store one or more rows associated with a database table in the allocated buffer;

10 communicate a message to one or more destination processing modules, the message comprising at least some of the one or more rows stored in the allocated buffer;  
otherwise:

execute a many-rows method to redistribute the one or more rows.

15 19. The computer-readable medium of Claim 18, wherein the transmitting processing module comprises one of a plurality of processing modules associated with a relational database system.

20 20. The computer-readable medium of Claim 18, wherein the message is communicated to each of a plurality of destination processing elements.

21. The computer-readable medium of Claim 18, wherein the allocated buffer is capable of storing no more than ten (10) rows.

22. The computer-readable medium of Claim 18, wherein the many-rows method comprises:  
25 communicating from one or more transmitting processing modules a first signal to a plurality of processing modules within a relational database system, the first signal operable to initiate a row receiver task on each of the processing modules;

communicating from one or more of the processing modules a ready-to-receive signal to the one or more transmitting processing modules;

30 communicating from the one or more transmitting processing modules a second signal comprising the one or more rows associated with the database table;

after communication of the last row associated with the database table, communicating from the one or more transmitting processing modules an end-of-data signal to each of the plurality of processing modules.

5        23. The computer-readable medium of Claim 18, further comprising invoking the program on a single transmitting processing module.

24. The computer-readable medium of Claim 18, further comprising determining a number of rows to store in the allocated buffer.

10

25. A relational database management system, comprising:
- a memory operable to store a program accessible to one or more of a plurality of transmitting modules, the program capable of managing a redistribution of one or more rows associated with one or more database tables; and
  - 5 at least one processor operable to determine the number of transmitting modules on which the program was invoked, the at least one processor capable of executing a few-rows row redistribution method to redistribute the one or more rows if the program was invoked on a single transmitting module, the at least one processor also capable of executing a many-rows row redistribution method to redistribute the one or more rows.

10

26. A database management system, comprising:
  - a massively parallel processing system comprising:
    - one or more nodes;
    - a plurality of processors, each of the one or more nodes providing access to one or more processors; and
      - a plurality of virtual processes, each of the one or more processors providing access to one or more virtual processes;
      - a set of one or more database tables residing on the one or more nodes, the one or more database tables containing information organized by geographic location; and
    - one or more of the plurality of virtual processes operable to:
      - allocate a buffer associated with a transmitting processing module, the transmitting processing module having access to a program, the program capable of managing a redistribution of one or more rows associated with one or more database tables;
      - compare the allocated buffer to a portion of the buffer to be occupied by the one or more rows;
      - if the allocated buffer is larger than the portion of the buffer to be occupied by the one or more rows:
        - store one or more rows associated with a database table in the allocated buffer;
        - communicate a message to one or more destination processing modules, the message comprising at least some of the one or more rows stored in the allocated buffer;
        - otherwise:
          - execute a many-rows method to redistribute the one or more rows.